DOI: ۱۰.۲۲۰۶۷\LJ.V۱۱I۲۱.۳۹۳۳۴

سنجشِ مهارت گوش‌کردن در محیط دانشگاه: بررسی موردی آزمون بسندگیِ فارسی در دانشگاه فردوسی مشهد

محسن رودمعجنی[۱]؛ (دانش‌آموخته دکتری زبان و ادبیات فارسی دانشگاه فردوسی مشهد، مشهد، ایران)
بهزاد قنسولی[۲]؛ (استاد گروه آموزشی زبان انگلیسی دانشگاه فردوسی مشهد، مشهد، ایران)
احسان قبول[۳]؛ (استادیار گروه آموزشی زبان و ادبیات فارسی، دانشگاه فردوسی مشهد، مشهد، ایران)

صص: ۲۵-۵۴

چکیده

اهمّیّت برخورداری از توانمندی گوش‌کردن در زندگی دانشگاهی دانشجویان قابل انکار نیست. افراد برای انجام بسیاری از فعّالیّت‌های دانشگاهی خود نیازمند برخورداری از این مهارت هستند. از همین رو، سنجش این توانمندی بخش مهمی از آزمون‌های بسندگی زبان فارسی را به خود اختصاص می‌دهد. از جمله آزمون زبان فارسی دانشگاه فردوسی مشهد که به‌عنوان معیاری برای سنجش دانشجویان غیرفارسی‌زبان برای ورود به دانشگاه‌های ایران، در برخی شهرهای ایران و عراق برگزار می‌گردد. در این پژوهش با کمک مدل آماری «راش»[۱]، به بررسی بخش گوش‌کردن این آزمون پرداختیم. نتایج نشان داد که پایایی آزمون ۰/۶۶ است که برای یک آزمون سرنوشت‌ساز میزان پایینی به‌حساب می‌آید. یکی از دلایل این ضعف را می‌توان تعداد کم سؤال و تکلیف‌های آزمون دانست. آنچه در آزمون زبان فارسی دانشگاه فردوسی مورد سنجش قرار می‌گیرد، تنها گوش‌کردن با هدف پیدا کردن اطّلاعات است و اهداف دیگر فعّالیّتهای گوش‌کردن در محیط دانشگاه مانند رسیدن به درک ابتدایی از متن، یادگیری و ادغام اطّلاعات در بین چند متن، نادیده مانده‌اند. همچنین نقشهٔ «آزمون‌دهنده–پرسش»[۲] بیانگر این نکته است که سؤالات آزمون

۱. RASCH
۲. Write Map

تاریخ دریافت :۱۳۹۷/۱۰/۲۱ تاریخ پذیرش :۱۳۹۸/۱۰/۱۶

پست الکترونیکی: ۱. mroodmajani@gmail.com ۲. ghonsooly@um.ac.ir ۳. ehsan.ghabol@yahoo.com

عمدتاً در سطح زبان‌آموزان متوسط و یا ضعیف هستند. برای بهبود پایایی می‌توان در طراحی آزمون از انواع دیگر متون مانند مشاوره‌ها و گفتگوها نیز استفاده کرد. از نظر نوع مطابقت، سؤالات آزمون عمدتاً از نوع «محل‌یابی» و یا «سلسله‌ای» هستند، قرار گرفتن تعداد بیشتری پرسش از نوع ترکیبی و تولیدی می‌تواند به سخت‌تر شدن سؤالات بیانجامد. به‌طورکلّی آزمون بسندگی زبان فارسی دانشگاه فردوسی بیشتر به سنجش توانمندی تشخیص واژگان کلیدی و درک جزئیات متن می‌پردازد. درک کارکرد ارتباطی پاره‌گفتارها، نتیجه‌گیری‌کردن برمبنای رابطه ی بین عناصر موجود در متن، پیوند دادن اطّلاعات یک متن شنیداری و یک متن خواندن از جمله توانمندی‌های دیگر است که استفاده از آنها می‌تواند سطح دشواری آزمون را ارتقاء دهد.

کلمات کلیدی: مهارت گوش‌کردن، آزمون بسندگی زبان، مدل راش، سنجش ارتباطی، پایایی آزمون.

مقدّمه

نمی‌توان اهمّیّت برخورداری از توانمندی گوش‌کردن[۱] در زندگی دانشگاهی دانشجویان را انکار کرد. بسیاری از فعّالیّت‌های دانشگاهی نیاز به تسلط به این مهارت دارد. درک سخنان استاد در کلاس درس، صحبت کردن با سایر هم‌کلاسی‌ها و دانشجویان، مشورت با استادان دربارهٔ موضوعات پژوهشی در خارج از کلاس، گفت‌وگو با کارمندان دانشگاه و حتّی سفارش غذا در سالن غذاخوری دانشگاه ازجمله فعّالیّت‌هایی هستند که انجام آنها در گرو بهره‌مندی از مهارت گوش‌کردن است.

داشتن این مهارت برای دانشجویان غیربومی زبان دانشگاه‌ها نیز از اهمّیّت بالایی برخوردار است و ناتوانی و ضعف در آن می‌تواند تأثیری جدّی بر موفقیت تحصیلی این دانشجویان داشته باشد. ازهمین‌رو است که بخشی قابل توجّهی از آزمون‌های بسندگی زبان دوم مانند «آیلتس»[۲] و «تافل»[۳]، به سنجش آن اختصاص دارد. دانشجویان غیرفارسی‌زبان دانشگاه‌های

۱. در این مقاله گوش‌کردن در معنای درک مطلب شنیداری به‌کارمی‌رود. از آنجایی‌که در گوش‌کردن میزانی از اراده و اختیار وجود دارد، عمل گوش‌کردن متفاوت از شنیدن است. شنیدن را می‌توان مقدّمه‌ای بر گوش‌کردن دانست.

۲. IELTS

۳. TOEFL

ایران نیز شرایطی مشابه دارند و اندازه‌گیری توانایی گوش‌کردن آنها می‌تواند به‌عنوان معیاری برای پیش‌بینی موفقیت تحصیلی آنها به‌شمارآید. از جمله آزمون‌هایی که در زبان فارسی به سنجش این مهارت می‌پردازد، آزمون بسندگی در مرکز بین‌المللی آموزش زبان فارسی دانشگاه فردوسی مشهد است که سالانه در دو نوبت در برخی شهرهای مختلف ایران و عراق مانند مشهد، اهواز، نجف و کوت برگزار می‌شود.[۱]

در این پژوهش، تلاش شده است تا به بررسی بخش گوش‌کردن این آزمون پرداخته و به دو سؤال دربارهٔ آن پاسخ داده شود. نخست اینکه این بخش از آزمون زبان فارسی فردوسی از چه میزان پایایی برخوردار است؟ دوم اینکه چه رابطه‌ای بین سطح سختی سؤالات این آزمون و سطح توانایی آزمون‌دهندگان وجود دارد؟ در حقیقت هدف پی‌بردن به این مسأله است که آیا آزمون از ثبات اندازه‌گیری برخوردار است و در دفعات مختلف سازه‌های یکسانی را می‌سنجد؟ و دیگر اینکه آیا سؤالات آزمون متناسب با سطوح زبانی مختلف آزمون‌دهندگان طراحی شده‌اند و آزمون، توانایی تمییز داوطلبان ضعیف، متوسط و قوی را دارد؟

به همین منظور نتایج به‌دست‌آمده از یکی از آزمون‌های برگزارشده در مرکز بین‌المللی زبان فارسی فردوسی را توسط مدل آماری «راش» مورد بررسی قرار دادیم. این مدل آماری که ساختاری پیچیده‌تر از مدل‌های کلاسیک دارد، از یک مزیّت اساسی برخوردار است و آن وابسته نبودن به نمونهٔ آماری پژوهش است (موسوی،۲۰۱۲؛ بکمن ۲۰۰۴:۱۳۹). مدل‌های آماری کلاسیک مانند «کودر ریچاردسون»[۲]، «روش دونیمه»[۳]، «آلفا کرونباخ»[۴] و «آزمون بازآزمون»[۵] بر این فرض اساسی استوار هستند که نمرهٔ به‌دست‌آمده از یک آزمون، حاصل جمع نمرهٔ حقیقی فرد و میزانی از خطا در محاسبه است. مبنی قرار دادن چنین فرضی باعث شده است که این مدل‌ها دارای دو ضعف اساسی باشند؛ نخست اینکه توانایی تمییز خطاهای

---

۱. هرچند طراحان این آزمون به طور رسمی مطلبی دربارهٔ آزمون خود چاپ نکرده‌اند، با این حال در تارنمای مرکز بین‌المللی آموزش زبان فارسی فردوسی نمونهٔ سوالات دوره‌های قبل این آزمون و نیز بخش‌های مختلف آن معرفی شده است. این آزمون سالانه برای سنجش توانایی زبانی حدود ۴۰۰ زبان آموز مورد استفاده قرار می‌گیرد و نتیجهٔ آن مورد پذیرش وزارت علوم می‌باشد.

۲. Kuder-Richardson

۳. Split half Experiment

۴. Cronbach’s Alpha

۵. Test-Retest experiment

نظام‌مند از خطاهای تصادفی را نداشته باشند (بکمن، ۱۹۹۰: ۱۸۶)، دوم اینکه با تغییر نمونهٔ پژوهش، نتایج به‌دست‌آمده نیز تغییر می‌کند.

به دلیل محدودیت‌های موجود در مدل‌های آماری کلاسیک، روان‌سنجان، مدل‌های آماری پیچیده‌تری را با عنوان «آی‌آرتی»[1] طراحی کردند. مدل‌های «آی‌آرتی» این قابلیت را دارند که عملکرد آزمون‌دهندگان در آزمون را با توجّه به توانمندی زبانی آنها مورد بررسی قرار دهند؛ همین امر باعث می‌شود که دیگر به نمونهٔ آماری پژوهش وابسته نباشند. مدل آماری «راش» نیز یک نمونه از مدل‌های «آی‌آرتی» به‌حساب می‌آید.

در این پژوهش سعی شده است تا سازه‌های تشکیل‌دهندهٔ مهارت گوش‌کردن و نیز ارتباط بین آنها مشخص شود. تعریف سازه‌های این مهارت، راه را برای سنجش دقیق و معتبر[2] توانایی گوش‌کردن هموار می‌کند. از سویی دیگر، چارچوب مشخصات تکلیف به‌عنوان روشی نوین و نظام‌مند از بررسی تکلیف‌های حقیقی زبانی[3] و شبیه‌سازی آنها در آزمون، برای اوّلین بار در زبان فارسی معرفی می‌شود. در پایان بر اساس مبانی نظری مطرح‌شده و نیز داده‌های به‌دست‌آمده از مدل آماری «راش»، تلاش شده است تا نقاط ضعف و قوّت آزمون زبان فارسی دانشگاه فردوسی بررسی شود.

پیشینهٔ پژوهش

شاید بتوان گفت که اوّلین شیوهٔ علمی سنجش مهارت گوش‌کردن به اوایل دههٔ ۶۰ قرن بیستم برمی‌گردد. در این دورهٔ آزمون‌سازی زبان تحت تأثیر دو جریان ساخت‌گرایی در زبان‌شناسی و رفتارگرایی در روان‌شناسی قرار داشت و باور بر این بود که می‌توان زبان را به عناصر کوچک‌تر سازندهٔ آن یعنی واج، تکیه، آهنگ، ساخت‌های نحوی و واژگانی تقسیم کرد و آنها را به‌صورت جداگانه در قالب سؤالات صحیح و غلط، سؤالات چندگزینه‌ای و تصاویر، سنجید (باک، ۲۰۰۱: ۶۲). امروزه از این رویکرد با نام «سنجش نکات مجزا»[4] یاد می‌کنند. این رویکرد، محصول عصر ساخت‌گرایی رواسنجی زبان بود و مبنایی فراهم آورد برای نیاز زبانی

---

۱. IRT

۲. Valid

۳. Real World Language Tasks

۴. Discrete point assessment

دورهٔ خود یعنی آزمون‌سازی صنعتی (اسپولسکی، ۲۰۰۸: ۴۴۹). در این دوره، شرکت‌ها و مؤسسات مختلف نیازمند برگزاری آزمون‌های زبانی در تعداد انبوه بودند؛ بنابراین باید امکان تصحیح آزمون با سرعت زیاد و تا حد امکان بدون دخالت نیروی انسانی فراهم می‌بود و سؤالات چهارگزینه‌ای بهترین نوع سؤال برای طراحان به شمار می‌آمد.

برای رویکرد «سنجش نکات مجزا» می‌توان یک مزیّت اساسی قائل بود و آن، قرار دادن اطّلاعات قابل شمارش در اختیار آزمون‌گران است (مورو، ۱۹۸۱: ۱۱). بااین‌حال، اگر به فلسفهٔ زبانی این رویکرد ایرادی وارد گردد کل بنای آن در هم خواهد ریخت، اتّفاقی که در دههٔ هفتاد رخ داد. در اوایل این دهه، پژوهشگرانی مانند اولر[۱] رویکرد دیگری را معرفی کردند که به «آزمون‌های یکپارچه»[۲] موسوم بودند. درحالی‌که نکات مجزا درپیِ سنجش دانش زبان به صورت اجزای منفصل و جداگانه بود، آزمون‌های یکپارچه توانایی کاربرد چندین عنصر سازندهٔ زبان را به‌صورت هم‌زمان موردسنجش قرار می‌دادند. اولر معتقد بود که تنها داشتن آگاهی از اجزای زبان کافی نیست، بلکه بین عناصر سازندهٔ زبان ارتباطاتی برقرار است که فرد باید از آنها نیز آگاه باشد؛ بنابراین اندازه‌گیری دانش فرد از اجزای مختلف زبان و جمع بستن آنها برابر با دانش زبانی او نیست، بلکه باید توانایی فرد در استفاده از چندین عنصر زبانی را به صورت هم‌زمان اندازه‌گیری کرد. ازجمله تکنیک‌های سنجش یکپارچه می‌توان به آزمون «تکمیل متن»، «دیکته» و «ترجمه» اشاره کرد (باک، ۲۰۰۱: ۶۲).

تقریباً هم زمان با اولر، شخص دیگری با نام کرول[۳] به نکته‌ای اشاره کرد که هرچند در زمان خود او چندان مورد توجّه قرار نگرفت، امّا با قدرت گرفتن جریان آموزش زبان ارتباطی، به رویکرد غالب در آزمون‌سازی زبان بدل گشت. کرول معتقد بود که آنچه اهمّیّت دارد، درک مطلوب پیام است؛ یعنی شنونده نه‌تنها باید اطّلاعات زبانی پیام را بفهمد، بلکه باید بتواند اطّلاعات درک‌شده را با بافت ارتباطی کلام پیوند دهد.

> هدف اصلی زبان، برقراری ارتباط در یک موقعیت خاص و با یک محیط خاص است. آنچه اهمّیّت دارد، دانش زبانی افراد و یا صحّت دستوری جملات

---

۱. Oller

۲. Integrated Tests

۳. Carroll

آنها نیست، بلکه توانایی آنها در استفاده از زبـان بـه‌منظـور برقـراری ارتبـاط در محیط زبانی مقصد است (باک، ۲۰۰۱: ۸۳).

بنابراین مفهومی که در رویکردهای «سنجش نکات مجزا» و «سنجش یکپارچه» بدان توجّـه نشده است، کاربرد زبان[۱] است. این دو رویکرد، طیف گسترده‌ای از اجزاء و ساختارهای زبـانی را می‌سنجند، اما توجّهی به توانایی به‌کارگیری زبان ندارند. تعامل‌محور بودن زبـان، غیـر قابـل پیش‌بینی بودن آن و تأثیر بافت و هدف در کاربرد زبان، از جمله ویژگـی‌هـایی هسـتند کـه در رویکردهای پیشین نادیده بوده‌اند. (مورو، ۱۹۸۱: ۱۶). این مسأله بدان می‌مانـد کـه شـما بـرای سنجش مهارت رانندگی یک فرد، سؤالاتی دربارهٔ کلاچ، دنده، سیستم برق ماشین و نیز علائـم رانندگی از او بپرسید بدون آنکه از او بخواهید یک ماشین را براند.

باوجود انتقادات تندی که طرفداران رویکـرد سـنجش ارتبـاطی بـر سـنجش نکـات مجـزا داشتند، آنها نیز به‌طور دقیق نمی‌توانسـتند مشـخص کننـد کـه چـه چیـزی را بایـد جـایگزین شیوه‌های قدیمی کرد. در حقیقت، سنجش ارتباطی در دوره‌های نخستین با دو مشـکل اساسـی روبه رو بود؛ نخست اینکه موقعیت‌هایی که افراد در آنها قرار می‌گیرنـد، معمـولاً بسـیار متنـوّع است و نمی‌توان تمام آن‌ها را در آزمون قرار داد و همچنین مشخص نیست کـه درک مطلـوب پیام که کرول از آن صحبت می‌کند، دقیقاً چیست؟ متـون محـل تفسـیرند و مـی‌تواننـد معـانی ضمنی مختلفی داشته باشند. چگونه می‌توان مناسب بودن یک پاسخ را مشـخص کـرد؟ (بـاک، ۲۰۰۱: ۶۲).

در سال ۱۹۹۰ بکمن با معرفی یک مفهوم تازه، مسیر جدیدی در جریـان سـنجش ارتبـاطی ایجاد کرد. این مفهوم، «اصالت تعامل»[۲] نام داشت. اصالت تعامل در حقیقت ارتباطی اسـت کـه بین آزمون‌دهنده و تکلیف امتحانی برقرار می‌شود. بکمن معتقـد اسـت کـه هـدف از سـنجش ارتباطی، تنها شبیه‌سازی فعّالیّتهای زبانیِ حقیقی در آزمون نیست بلکه آزمون‌گران بایـد بتواننـد با شناسایی عوامل تأثیرگذار در عملکرد آزمون‌دهندگان، بخش‌های مناسـبی از کـاربرد حقیقـی زبان را شبیه‌سازی کنند (بکمن،۱۹۹۰: ۳۱۷)؛ بنابراین بر خلاف باورهای نخسـتینِ سنجشـگران ارتباطی، آزمونی که سؤالات آن شبیه به فعّالیّتهای زبانی حقیقی اسـت، لزومـاً آزمـون معتبـری

---

۱. Language Use

۲. Interactional Authenticity

نیست؛ بلکه آزمون باید ویژگی‌هایی از کاربردِ حقیقی زبان را دربرداشته باشد که با تفاسیر برآمده از نمرات آزمون و همچنین کاربرد آن نمرات، متناسب باشد. به‌منظور بررسی نظام‌مند تکلیف‌های[1] زبانی، بکمن چارچوبی نظری ارائه می‌دهد که در بخش مبانی نظری این پژوهش به‌تفصیل به بررسی نمونه‌ای مشابه آن با عنوان «چارچوب مشخصات تکلیف»[2] پرداخته خواهد شد. تردیدی نیست که امروزه سنجش ارتباطی به رویکرد مسلط در عالَم آزمون‌سازی بدل شده است (هاردینگ، ۲۰۱۴: ۱۸۸). مجموعه آزمون‌های دانشگاه کمبریج، آزمون «آیلتس» و آزمون نوین «تافل» از نمونه آزمون‌هایی هستند که بر مبنای اصول ارتباطی طراحی شده‌اند. همچنین طراحان مجموعه‌های تأثیرگذاری مانند «چارچوب مرجع اروپا» که مبنای طراحی کتاب‌های آموزشی بسیار قرار گرفته است، باورهای ارتباطی دارند.

لازم به یادآوری است که در زبان فارسی نیز تلاش‌هایی برای طرّاحی آزمون بسندگی صورت گرفته است (موسوی، ۱۳۷۹؛ قنسولی، ۱۳۸۹؛ جلیلی، ۱۳۹۰؛ گلپور، ۱۳۹۴) و هر یک از این پژوهش‌ها بخشی از آزمون خود را به مهارت گوش‌کردن اختصاص داده‌اند. مشخصات این آزمون‌ها در جدول شمارهٔ (۱) آمده است. همان‌طور که دیده می‌شود هیچ‌یک از این پژوهش‌ها، تعریف دقیقی از سازه‌ها و فرآیندهای مهارت گوش‌کردن و نیز مشخصات تکالیف زبانی ارائه نداده‌اند. همچنین در تمام این پژوهش‌ها، مدل‌های آماری کلاسیک به کار رفته است. همان‌طور که پیش‌تر گفته شد، این مدل‌های آماری به نمونهٔ آزمون، وابسته بوده و با تغییر نمونه، نتایج آنها نیز دچار تغییر می‌شود.

مبانی نظری

رسیدن به یک سنجش دقیق از مهارتِ گوش‌کردن در محیط دانشگاه، نیازمند پاسخ‌گویی به دو پرسش اساسی است؛ نخست اینکه سازه‌های تشکیل‌دهندهٔ مهارت گوش‌کردن چیست؟ و این سازه‌ها چه ارتباطی با یکدیگر دارند؟ دوم اینکه دانشجویان در محیط دانشگاهی با چه تکلیف‌هایی روبه‌رو می‌شوند؟ و این تکلیف‌ها از چه ویژگی‌هایی برخوردارند؟ برای رسیدن به پاسخ پرسش نخست، در ابتدا به بررسی سازه‌های تشکیل‌دهندهٔ مهارت گوش‌کردن و نیز

---

۱. تکلیف زبانی به هر نوع فعّالیّتی گفته می‌شود که موقعیتی مشخص داشته و شخص برای رسیدن به اهدافی از پیش تعیین شده به آن می‌پردازد (کرول،۱۹۹۲: ۸).

۲. Language task charactristics

ارتباط این سازه‌ها با یکدیگر پرداخته شد. برای پرسش دوم نیز همان‌طور که پیش‌تر گفته شد، چارچوب مشخصات تکلیف آزمون «تافل» معرفی خواهند شد.

جدول (۱): آزمون‌های طراحی‌شده در زبان فارسی برای سنجش مهارت گوش‌کردن

نویسنده	تاریخ	هدف از انجام پژوهش	روش آماری بررسی پایایی پژوهش	سازه‌ها و فرآیندهای مهارت گوش کردن	ویژگی‌های فعّالیّت‌ها یا تکلیف‌های زبانی
موسوی	۱۳۷۹	طراحی آزمون بسندگی زبان فارسی برای سنجش دانش زبانی غیرفارسی‌زبانان	روش‌های آماری کلاسیک	تعریفی ارائه نشده است	نامشخص است
قنسولی	۱۳۸۹	طراحی آزمون بسندگی زبان فارسی	آلفا کرونباخ	تعریفی ارائه نشده است	نامشخص است
جلیلی	۱۳۹۰	طراحی آزمون بسندگی زبان فارسی برای سنجش دانش زبانی غیرفارسی‌زبانان	بدون تحلیل آماری	تعریفی ارائه نشده است	نامشخص است
گلپور	۱۳۹۴	طراحی آزمون بسندگی زبان فارسی	روش‌های آماری کلاسیک	تعریفی ارائه نشده است	نامشخص است

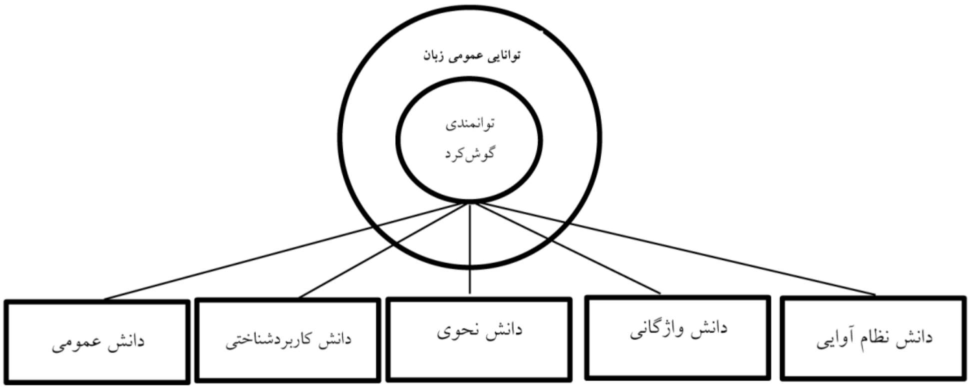
سازه‌های تشکیل‌دهندهٔ مهارت گوش‌کردن

آگاهی از سازه‌های تشکیل‌دهندهٔ مهارت گوش‌کردن برای رسیدن به یک سنجش دقیق از این مهارت، الزامی است. مایکل راست[1] معتقد است که توانایی گوش‌کردن بخشی از توانایی عمومی زبانی فرد است و می‌توان آن را به پنج جزء کوچک‌تر تقسیم کرد که عبارت اند از: ۱ـ

۱. Michael Rost

دانش نظام آوایی[۱] ۲ـ دانش واژگانی[۲] ۳ـ دانش نحوی[۳] ۴ـ دانش کاربردشناختی و [۴] ۵ـ دانش عمومی[۵] (راست، ۲۰۱۶: ۱۹۷). در ادامه این سازه‌ها از دیدگاهِ راست معرفی خواهند شد.

شکل شمارهٔ (۱): توانایی عمومی زبان و توانمندی شنیدن (همان).



برجسته‌ترین ویژگی گوش‌کردن که آن را از سایر مهارت‌ها متمایز می‌کند، نیاز به آگاهی از نظام آوایی زبان است. توانایی تشخیص واج‌ها و گروه‌های واجی، توانایی تشخیص تغییرات واج‌گونه‌ها در گفتار سریع -به‌عنوان‌مثال، می‌خورن به‌جای می‌خورند- آگاهی از واحدهای زبرزنجیری گفتار مانند تکیه، آهنگِ کلام و تشخیص صورت گفتاری واژگان ازجمله دانش‌های نظام آوایی زبان هستند که آگاهی از آن‌ها برای درک زبان گفتار ضروری است.

بااین حال یکی از اساسی‌ترین واحدهای زبانی که در درک زبان گفتار اهمّیّت بسیاری دارد، واژگان است. تشخیص خودکار واژگان در گفتار سریع و روان از جنبه‌های بسیار حیاتی فراگیری زبان اوّل و دوم به‌شمار می‌آید. هنگامی‌که شنوندگان با مشکلاتی در درک اصوات روبه رو می‌شوند و صداها برایشان واضح نیست ـ مخصوصاً اگر ساختارهای دستوری نیز نامفهوم باشند ـ عموماً بر اطّلاعات واژگانی خود تکیه می‌کنند. امّا تشخیص واژگان، خود، از چهار توانمندی تشکیل شده است: ۱ـ تشخیص ساخت‌های واژگانی پایه؛ به‌عنوان‌مثال برای

۱. Phonological Knowledge
۲. Lexical Knowledge
۳. Syntacitic knowledge
۴. Pragmatic knowledge
۵. General knowledge

پی‌بردن به معنای واژهٔ ورزش‌کار، شنونده باید ساخت پایهٔ این ترکیب، یعنی ورزش را بداند ۲ـ آگاهی از معنای واژه ۳ـ دانستن این نکته که واژه معمولاً در کنار چه واژگان دیگری به کار می‌رود؟ و ۴ـ آگاهی صورت‌های مجاز واژگانی (ورزشکار).

تا هنگامی که اطّلاعات به‌دست‌آمده از واژگان درون یک مدل دستوری سازمان‌دهی نشوند، با مجموعه‌ای درهم‌ریخته از اطّلاعات روبه‌رو هستیم. دانش موجود در ذهن یک فرد این امکان را برای او فراهم می‌آورد که بتواند زبان را در سطوحی فراتر از واژگان، یعنی جمله و گفتمان، تحلیل کند. دانش نحوی، خود، شامل توانایی تجزیهٔ نحوی در سطح جمله، توانایی تجزیهٔ نحوی در سطح گفتمان، تشخیص همایندها و تشخیص نشانه‌های انسجامی در گفتمان است.

همان‌طور که مشاهده کردیم، اطّلاعات شنونده دربارهٔ نظام‌های آوایی زبان و ساختارهای واژگانی و نحوی در درک زبان گفتار از اهمّیّت بسیار بالایی برخوردار هستند. بااین‌حال این اطّلاعات زمانی سودمند خواهند بود که ما بتوانیم آنها را با بافت و موقعیتی که زبان در آن اتّفاق می‌افتد، پیوند دهیم. این بدان معناست که ما علاوه بر اطّلاعات بالا، به دانش کاربردشناختی زبان نیز نیازمندیم. بررسی کاربردِ شناختی زبان، یعنی توصیف یک پدیدهٔ زبانی از منظر شخصی گوینده و شنونده در امر بینافردی که در یک تعامل دوسویه ایجاد می‌شود. این حضور شنونده در تعامل زبانی، نیازمند مشارکت و تعامل است و این مشارکت باید متناسب با نقش شنونده باشد. به‌عنوان‌مثال، با توجّه به اینکه اطّلاعات تازه در متن معمولاً آهنگ افتان دارند و اطّلاعات ازپیش‌داده‌شده آهنگ خیزان، شنونده باید بتواند جریان اطّلاعات را دنبال کند. او همچنین باید بتواند به کمک دانش خود، به کنش‌های گفتاری زبان منظور و نیز قصد گوینده پی ببرد. همچنین فرد باید بتواند از تجربیات خود در مواجهه با متون و موقعیت‌های مشابه استفاده کند و به کمک آنها به تفسیر موقعیت حاضر بپردازد.

توانمندی گوش‌کردن از مجموعه‌ای از دانش‌های دیگر نیز تشکیل شده که مایکل راست از آنها با نام دانش عمومی یاد می‌کند. درک یک متن شنیداری زمانی میسّر خواهد شد که فرد از ارتباطات بین افراد و نیز موضوعات عمومی اطّلاعاتی داشته باشد. همچنین آگاهی از علوم مختلف، مانند تاریخ، جغرافیا، علوم تجربی و ریاضی و نیز آشنایی با راهکارهای استفاده از دانش از جمله توانمندی‌هایی است که در بخش دانش عمومی قرار می‌گیرد.

مدلی برای مهارت گوش‌کردن

پس از مشخص کردن سازه‌های مهارت گوش‌کردن، پرسش دیگری مطرح می‌شود و آن چگونگی ارتباط این سازه‌ها با یکدیگر است. بجار و همکاران تصویری از فرآیندهای تشکیل‌دهندهٔ مهارت گوش کردن ارائه داده‌اند که در شکل شمارهٔ دو آمده است (۲۰۰۰: ۳). در وهله‌ی نخست، درونداد صوتی توسط سیستم شنیداری دریافت شده و تبدیل به پیام‌هایی قابل درک برای مغز انسان می‌شود. سپس فرآیندهای شناختی با به‌کارگیری دانش‌های زبانی، موقعیتی و نیز دانش عمومی به پردازش پیام‌های دریافت شده می‌پردازند. محصول پایانی این بخش، درک مطلب شنیداری از درونداد نام دارد. از آنجایی که دانش‌های زبانی، موقعیتی و عمومی افراد با یکدیگر تفاوت دارد، بنابراین درک شنیداری آنها از یک درونداد صوتی یکسان، متفاوت است.

منظور از دانش زبانی، اطّلاعات فرد از آواهای زبان، واژگان، ساختارهای نحوی، گفتمان و دانش کاربردشناختی است. دانش موقعیتی نیز به نقش بافت[۱] در فعّالیّت گوش‌کردن اشاره دارد که شامل اطّلاعاتی می‌شود که فرد از شرکت‌کنندگان در فعّالیّت زبانی دارد و یا سرنخ‌های بصَری که به او در درک پیام کمک می‌کنند. دانش پیشینی نیز محصول تجربیات پیشین او در مواجهه با فعّالیّت‌های مشابه است. فرآیندهای شناختی شامل راهبردهای حل مسأله، بازیافت اطّلاعات پیشین و تفکر می‌شود.

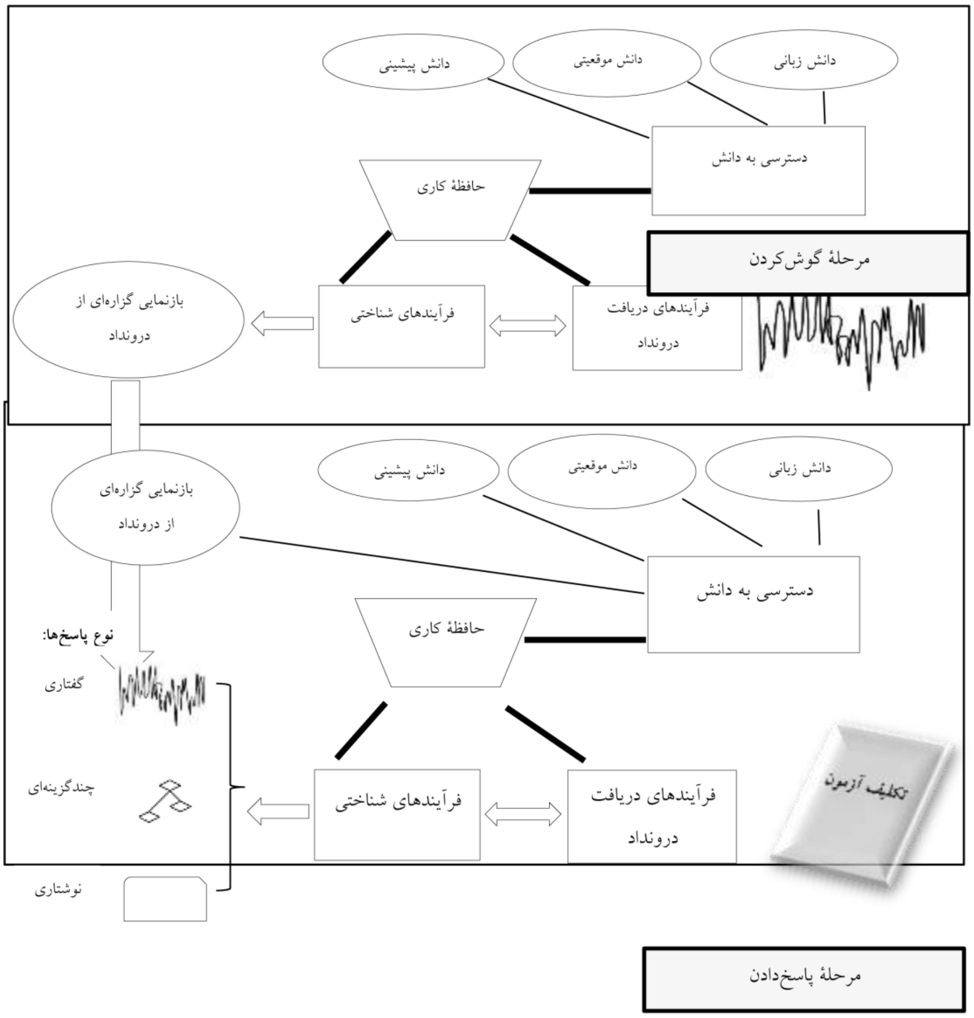
باید این نکته را در نظر داشت که سنجش مهارت گوش‌کردن به‌طور مستقل ممکن نیست و برای اندازه‌گیری آن باید فرد به سؤال و یا مجموعه‌ای از سؤالات پاسخ دهد؛ بنابراین سنجش این مهارت از دو بخش تشکیل می‌شود. همان‌طور که توضیح داده شد، بخش اوّل «درک شنیداری» نام دارد و بخش دوم «پاسخ‌دهی»[۲]. در بخش پاسخ‌دهی فرد به کمک مطالب درک‌شده از بخش شنیداری و با به‌کارگیری دانش‌های زبانی، عمومی و موقعیتی به سؤالات آزمون پاسخ می‌دهد. محصول پایانی این بخش می‌تواند به‌صورت گفتاری چندگزینه‌ای و یا نوشتاری باشد.

---

۱. Context

۲. Responce

شکل (۲): فرآیندهای گوش‌کردن و پاسخ‌دهی (بجار و همکاران،۲۰۰۰: ۳)



دانش‌هایی که در مدل بجار و همکاران (۲۰۰۰) به‌عنوان بخشی از فرآیند درک مطلب شنیداری دیده می‌شود، شباهت بسیاری با دانش‌های معرفی‌شده توسط مایکل راست دارد. منظور از دانش زبانی در مدل بجار و همکاران، همان دانش‌های آوایی، واژگانی، نحوی و کاربردشناختیِ راست است. دانش موقعیتی و پیشینی نیز در تصویر بالا، می‌تواند همان دانش عمومی در مدل راست باشد. بااین‌حال، مدل بجار و همکاران این امکان را به ما می‌دهد که علاوه بر اطّلاع از سازه‌های تشکیل‌دهندهٔ مهارت گوش‌کردن، از نحوهٔ ارتباط آنها با یکدیگر

نیز آگاه باشیم. در حقیقت این مدل، تلاشی است برای شناخت مراحلی که امواج صوتی را به پاسخ‌های داوطلبان در جلسهٔ آزمون تبدیل می‌کند.

چارچوب مشخصات تکلیف

همان‌طور که پیش‌تر گفته شد، شیوهٔ دیگر طرّاحی آزمون بررسی تکلیف‌های زبانی است که افراد در محیط موردنظر با آنها روبه‌رو می‌شوند. به این منظور، جمیسون[۱] و همکاران (۲۰۰۰) چارچوبی برای بررسی تکلیف‌های زبانی با نام «چارچوب مشخصات تکلیف» طرّاحی کرده‌اند. چارچوب مشخصات تکلیف شیوه‌ای کارآمد برای بررسی مشخصات فعّالیّتهای زبانی در محیط مقصد ارائه می‌دهد. استفاده از این چارچوب می‌تواند به اصالت آزمون کمک کند و محصول آن، درواقع، ارائهٔ تصویری واقعی‌تر از توانمندی آزمون‌دهندگان است. این چارچوب از سه بخش اساسی تشکیل شده است: ۱ـ موقعیت ۲ـ محتوای متن ۳ـ مشخصات آزمون (۲۰۰۰: ۱۳).

موقعیت

موقعیت دربرگیرندهٔ عناصری فرا زبان‌شناسی است که به فعّالیّتهای زبانی مربوط می‌شوند. برای مشخّص‌کردن موقعیت تکلیف باید به چهار پرسش پاسخ داد؟ ۱ـ چه کسانی در انجام تکلیف شرکت دارند؟ ۲ـ محتوای تکلیف چیست؟ ۳ـ محیطی که تکلیف در آن اتّفاق می‌افتد کجاست؟ ۴ـ هدف از انجام تکلیف چه است؟ (۲۰۰۰: ۱۴).

شرکت‌کنندگان

برخی از نکات مربوط به شرکت‌کنندگان مانند سن، نژاد و جایگاه طرف‌های صحبت از نظر قدرت، باعث تغییر در سختی تکلیف می‌شود؛ بنابراین باید این نکات را در طرّاحی آزمون در نظر داشته باشیم. پاورز (۱۹۸۶) بر اساس پژوهشی که در ۳۴ مؤسسه در شش رشته (مهندسی، روان‌شناسی، انگلیسی، شیمی، علوم کامپیوتر و تجارت) انجام داد، به این نتیجه رسید که طرف‌های صحبت با دانشجویان بین‌المللی هفت دسته هستند (۱۹۸۶: ۳۰–۳۱):

---

۱. Jamieson

جدول (۲): چارچوب مشخصات تکلیف (جمیسون و همکاران،۲۰۰۸: ۶۴)

مشخصات تکلیف	متغیّرهای تکلیف	تعریف متغیّرها
موقعیت	شرکت‌کنندگان	افرادی که در فعّالیّت زبانی شرکت دارند (مانند دانشجویان و استادان). می‌توان آنها را از نظر جنسیت، نژاد، سن و همچنین نقش اجتماعی دسته‌بندی کرد.
	محتوا	محتوا همان موضوع تکلیف‌های زبانی است. ـ تخصّصی ـ دانشگاهی که می‌بایست در حوزه‌های علوم اجتماعی و طبیعی باشند؛ زیرا گسترده هستند و موضوعات عمومی‌تری را دربر می‌گیرند؛ ـ کلاس درس: مربوط به فعّالیّت‌ها و تمرین‌های کلاسی؛ ـ غیردانشگاهی: موضوعات پرکاربردی که لزوماً به فضای کلاسی مربوط نمی‌شوند (بیماری، آب‌وهوا و...).
	محیط	فضایی که تکلیف در آن اتّفاق می‌افتد: ـ آموزشی، مانند کلاس درس آزمایشگاه و... ـ دانشگاهی ـ غیر آموزشی، مانند سالن مطالعه، خوابگاه و... ـ اداری، مانند درمانگاه، سالن غذاخوری و...
	هدف	هدف، شرکت در فعّالیّت زبانی یا همان کاربرد تکلیف ـ به‌عنوان‌مثال آموزشی ـ بیان احساسات فردی، برقراری تعامل
	گونه[۱]	میزان رسمی بودن متن که شامل سه سطح است: ـ رسمی مانند سخنرانی، ارائهٔ کلاسی، مقاله ـ نیمه‌رسمی: گفت‌وگو در اتاق استاد، گفت‌وگوی تلفنی با یک فرد ناشناس ـ غیررسمی
محتوای متن	ویژگی‌های دستوری	ساختارهای نحوی و واژگانی متون که می‌توان آنها را با استفاده از فرمول‌های خوانایی، بررسی بسامد واژگان، نوع افعال و بندهای وابسته بررسی کرد.

۱. Register

	ویژگی‌های کاربردی[۱]	ازآنجایی‌که عمدهٔ متون دانشگاهی تشریحی هستند، کاربرد اصلی این متون انتقال و یا کاویدن اطّلاعات است. کارکردهای دیگر می‌توانند دفاع کردن از یک ایده و یا ارائهٔ یک تحلیل (بحث، جدل) باشند. تعداد محدودی از متون تشریحی هدف اقناعی و یا تعاملی دارند.
	ویژگی‌های گفتمان	هدف نویسنده و یا گوینده که می‌تواند معنی کردن، توصیف کردن، بیان علّت و معلول، بیان مسأله و راه‌حل آن، آوردن مثال و نمونه، مقایسه کردن، ارائهٔ یک شرح ساده مانند فهرست اقلام و یا بیان قوانین و مقرّرات باشد.
مشخصات آزمون	سؤالات	نوع اطّلاعات خواسته‌شده: هرچه اطّلاعات خواسته‌شده انتزاعی‌تر باشد، سؤال سخت‌تر خواهد بود (به پیوست ۱ مراجعه کنید).
		نوع مطابقت: برای پاسخ‌گویی به سؤال، باید با چه روشی متن را بررسی کرد (به پیوست ۲ مراجعه کنید). ـ محل‌یابی: پیداکرده محل جواب در متن؛ ـ سلسله‌ای: پیداکرده چند محل جواب در متن؛ ـ ترکیبی: از دو مرحله تشکیل می‌شوند: نخست چند محل جواب در متن شناسایی شود، سپس رابطه‌ی بین آنها مشخص گردد؛ ـ تولیدی: سؤالاتی که در آن داوطلبان باید برای پیداکرده محل پاسخ به تفسیر بپردازند.
		حضور گزینه‌های انحرافی: ساده‌ترین سؤالات سؤالاتی هستند که گزینهٔ انحرافی ندارند. هرچه تعداد گزینه‌های انحرافی بیشتر باشد و گزینه-های انحرافی اشتراکات بیشتری با پاسخ داشته باشند، سؤال سخت‌تر می‌شود.
	نوع پاسخ‌دهی	سؤالات چندگزینه‌ای، سؤالات تشریحی، کامل کردن یک نمودار.
	قواعد نمره-دهی	مراحل نمره‌دهی به پاسخ‌های آزمون‌دهندگان.

۱- استاد در کلاس درس؛

۱. Pragmatic Features

۲- کتابدار (برای مثال، در حالِ دادنِ اطّلاعات دربارۀ سیستم جدید کتابخانه)؛
۳- مسئول ثبت‌نام و یا منشی گروه که توضیحاتی دربارۀ مراحل ثبت‌نام می‌دهد؛
۴- راهنمای دانشجویان بین‌الملل؛
۵- مسئول آزمایشگاه و یا دستیار استاد که به دانشجویان دربارۀ انجام یک آزمایش و یا تکالیف توضیح می‌دهد؛
۶- پرستار و یا دکتر در مرکز درمانی دانشگاه؛
۷- هم‌کلاسی که دربارۀ یک امتحان پیشنهادهایی می‌دهد.

هرچند این دسته‌بندی برمبنای پژوهش صورت گرفته در دانشگاه‌های آمریکای شمالی است، امّا به دلیل نبود پژوهش مشابه در زبان فارسی، از آن برای مشخّص‌کردن شرکت‌کنندگان در تکالیف مهارت گوش‌کردن استفاده کردیم.

محیط

یک گفت‌وگوی علمی را در نظر بگیرید که بین تعدادی از دانشجویان در داخل سالن غذاخوری دانشگاه در جریان است و حالا تصوّر کنید همان گفت‌وگوی علمی در یک جلسۀ رسمی صورت بگیرد. نحوۀ مواجهه ی افراد با این دو متفاوت خواهد بود و این تفاوت نه در متن گفت وگو، بلکه تحت تأثیر محیط است؛ بنابراین باید محیط‌های مختلفی را که فعّالیّتهای مهارت گوش‌کردن در آن‌ها اتّفاق می‌افتد، شناسایی کنیم. بجار (۲۰۰۰) محیط را در فعّالیّتهای زبانی دانشگاهی این‌گونه تعریف می‌کند: ۱ـ محیط آموزشی (سالن سخنرانی، کلاس درس، آزمایشگاه، سالن سمینار) ۲ـ محیط مطالعه (سالن مطالعه ی خوابگاه، کتابخانه، اتاق استاد و مرکز کامپیوتر) ۳ـ محیط اداری (مرکز درمانی، کتاب‌فروشی، اتاق ثبت‌نام، اتاق مسئول دانشجویان بین‌المللی، امور مالی) (۲۰۰۰: ۷).

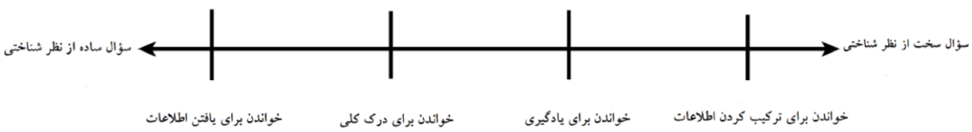
محتوا و هدف

محتوای آزمون همان موضوع فعّالیّت زبانی است که می‌بایست به‌اندازۀ کافی عمومی باشد تا اطّلاعات آزمون‌دهندگان در یک حوزۀ خاص باعث عملکرد بهتر آنها نشود. طرّاحان «تافل» موضوعات بخش گوش‌کردن را به سه دسته تقسیم می‌کنند: موضوعات دانشگاهی (علوم تجربی، علوم اجتماعی، علوم و هنر، علوم فیزیک)، موضوعات مربوط به کلاس (تکالیف،

زمان انجام تکالیف، کتاب‌های مربـوط بـه کـلاس) و موضـوعات مربـوط بـه فضـای پـردیس دانشگاه (ثبت‌نام، درمان، کتابخانه، راهنمای امور دانشگاه).

طرّاحان «تافل» همچنین چهار هدف برای گوش‌کردن قائل هستند: ۱ـ گوش‌کردن به‌منظور پیدا کردن اطّلاعات ۲ـ گوش‌کردن برای رسیدن به درک ابتدایی از متن ۳ـ گوش‌کردن با هدف یادگیری ۴ـ گوش‌کردن برای ادغام اطّلاعات در بین چند متن مختلف. این اهداف دارای سطوح سختی متفاوتی هستند که در شکل شمارهٔ ۳ نشان داده شده‌اند.

شکل (۳): مقیاس مفهومی فرضی برای اهداف خواندن (جمیسون و همکاران، ۲۰۰۸: ۵۵).



متن تکلیف

در بخش محتوای متن، ویژگی‌های متنیِ[1] تکلیف‌ها بررسی می‌شود. این بخش شـامل یـک پرسش است؛ متن تکلیف‌های زبانی در محیط مقصد دارای چه ویژگی‌های نحوی، واژگـانی و آواشناسی هستند؟

برای بررسی پیچیدگی ساختار نحوی زبان گفتار در زبان انگلیسی می‌توان به پژوهش بایبر[2] (۱۹۹۵) رجوع کرد، امّا در زبان فارسی تحقیقـی دربـارهٔ سـطح سـختی سـاختارهای دسـتوری صورت نگرفته است. تنها پژوهشی که به این حوزه مربوط می‌شود، «تعیین بسامد زمـان‌هـا در زبان گفتار» است که توسّط وکیلی‌فرد و همکاران (۱۳۹۲) انجام شده است.

واژگان کم‌بسامد نیز ازجمله عوامل سخت‌کنندهٔ زبـان گفتـار بـه شـمار مـی‌آینـد (بجـار و همکاران،۱۵:۲۰۰). برای مشخص کردن بسامد واژگان می‌توان از *فرهنگ بسامدی زبان فارسـی* راتلج کمک گرفت. در این فرهنگ، ۵۰۰۰ واژهٔ پربسامد زبان فارسی بر اساس یک پیکـرهٔ ۱۵۰ میلیون واژه‌ای مشخص شده است (میلر و همکاران، ۲۰۱۸).

---

۱. Textual

۲. Biber

بخش دیگری از چهارچوب مشخصات تکلیف به ویژگی‌های آواشناسی متن می‌پردازد، مانند «سرعت گفتار»، «پدیده ی سندی»[۱] و «مکث». میانگین سرعت گفتار در زبان فارسی ۳/۹۳ هجا در ثانیه است (مخلصین، ۱۳۹۶: ۳۳۰)؛ بنابراین هرچه سرعت گفتار از این میزان بیشتر باشد، تکلیف سخت‌تر از میزان معمول می‌شود. همچنین باور بر این است که مکث‌ها باعث ساده‌تر شدن گفتار برای شنونده می‌شوند. متداول‌ترین نوع مکث‌ها در زبان فارسی، واژه‌های مکث‌پرکنِ «خب»، «چی میگن» و حروف کشیده‌ی «بععععد» و «حالاااااا» هستند. شروع‌های اشتباهی ردیابی‌شده و نیز تکرار واجی «مmm من» نیز از بسامد بالایی برخوردار هستند (عرفانیان قونسولی و همکاران، ۱۳۹۳: ۱۱۰).

علاوه بر این‌ها، زبان گفتار دارای ویژگی‌هایی است که در زبان‌شناسی به «پدیده ی سندی» معروف‌اند، مانند «همسانگردی» و «حذف». نبودن این مشخصات طبیعتاً باعث ساده‌ترشدن زبان می‌شود، امّا میزان تأثیرگذاری آنها در سختی متن مشخص نیست. بااین‌حال توصیّه می‌شود به‌عنوان مشخّصه‌ی برجسته‌ی زبان گفتار، در طرّاحی تکالیف امتحانی در نظر گرفته شوند.

مشخصات آزمون

تاکنون به بررسی ویژگی‌های فعّالیّتهای زبانی در محیط دانشگاه پرداختیم که می‌توانند در سنجش مهارت گوش‌کردن تأثیرگذار باشند. امّا اگر بخواهیم این تکلیف‌ها را به سؤالات امتحانی تبدیل کنیم، می‌بایست چند نکته را روشن نماییم که بخش پایانی چارچوب انرایت با نام مشخصات آزمون را تشکیل می‌دهند؛ مباحثی از قبیل پرسش‌های آزمون دارای چه مشخّصاتی هستند؟ نوع سؤالات آزمون چیست؟

پرسش‌ها تأثیر بسیار مهمی در سختی آزمون دارند و ما می‌بایست میزان سختی آنها را به دقّت محاسبه نماییم. به همین منظور، جمیسون و همکاران (۲۰۰۰) به بررسی ویژگی‌های پرسش‌ها و تأثیر آنها در سختی آزمون پرداخته‌اند که آگاهی از این نکات می‌تواند در طرّاحی آزمون بسیار تأثیرگذار باشد. پرسش‌ها را می‌توان از دو منظرِ نوع اطّلاعات خواسته‌شده و نوع رابطهٔ سؤال و پاسخ بررسی کرد. هرچه اطّلاعات خواسته‌شده انتزاعی‌تر باشند، سؤال نیز از

---

۱. Sandhi Phenamenon

میزان سختی بیشتری برخوردار خواهند بود. رابطهٔ بین نوع اطّلاعات و سختی سؤال در جدول (۳) به‌طور کامل آمده است.

مطلب دیگری که در طرّاحی سؤال باید بررسی شود، رابطهٔ بین سؤال و پاسخ است. آیا داوطلب تنها باید محل یک پاسخ را در متن پیدا کند (محل‌یابی) و یا پاسخ در چند نقطه از متن قرار گرفته است و داوطلب می‌بایست محل‌های مختلفی را شناسایی کند (سلسله‌ای). در برخی از سؤالات، داوطلب علاوه بر پیدا کردن چند محل پاسخ در متن، باید رابطهٔ آنها را نیز با یکدیگر بررسی کند (ترکیبی) و یا در مواردی برای یافتن پاسخ به تفسیر بپردازد (تولیدی). میزان سختی هرکدام از این نوع سؤالات در جدول (۴) آمده است.

جدول (۳): نوع اطّلاعات خواسته‌شده (جمیسون و همکاران، ۲۰۰۰:۶۵-۶۳)

سطح سختی	نوع اطّلاعات خواسته‌شده
۱	تشخیص یک شخص، گروه، مکان یا شیء.
۲	پرسش‌هایی دربارهٔ زمان، مقدار، صفت، نوع، فعّالیّت، محل، موقعیت، گروه و یا مراحل انجام یک کار.
۳	پرسش‌هایی دربارهٔ حالت، برنامهٔ یک شخص، هدف و یا کارکرد، گزینهٔ انتخابی (مثال: نویسنده چه گزینه‌هایی برای کاهش آلودگی ارائه می‌دهد؟)؛ اقدام (او برای رئیس‌جمهور شدن چه کارهایی انجام داد؟)؛ وضعیت، مراحل، ارجاع ضمیر، تأیید (آیا درست است که مارکوپُلو دو بار به چین سفر کرد؟)؛ بیان یک مسأله؛ بیان راه‌حل.
۴	سؤالاتی که برای پاسخ دادن به آن‌ها می‌بایست رابطهٔ علّت و معلولی و شباهت چند چیز را بررسی کرد و یا سؤالاتی دربارهٔ نظر و عقیده، بیان الگو و یا تبیین یک موضوع.
۵	سؤالاتی مربوط به موضوع یک متن یا بند، بررسی تفاوت‌ها و یا معادل‌یابی برای واژگان کم‌بسامد.

مدلی برای سنجش مهارت گوش‌کردن در محیط دانشگاه

بر اساس مطالبی که دربارهٔ سازه‌های آزمون و تکلیف‌های زبانی در محیط دانشگاه گفته شد، می‌توان مدلی نهایی برای سنجش مهارت گوش‌کردن ارائه داد. این مدل در جدول (۵) معرفی شده و مبنای طرّاحی آزمون جدید «تافل» است. برای سنجش مهارت گوش‌کردن در

محیط دانشگاه هفت توانمندی باید در نظر گرفته شوند. سه توانایی نخستین، توانایی درک جزئیات متن، درک ایده‌ها، موضوع‌ها و اهداف اصلی و نیز درک واژگان کلیدی متن را شامل می‌شود که می‌توان آنها را به‌صورت پرسش‌های مربوط به چه کسی، چه چیزی، چه زمانی، چه مکانی، تشخیص و یا توضیح ایده‌های اصلی متن و تشخیص و درک واژگان کلیدی سنجید. این نوع سؤالات در حقیقت مربوط به سازه‌های آوایی، واژگانی و نحوی درک شنیداری می‌شوند و با اهداف گوش‌کردن برای پیدا کردن اطّلاعات و گوش‌کردن برای یادگیری مطابقت دارند.

جدول (۴): نوع مطابقت (همان:۶۶)

سطح سختی	نوع مطابقت
۱	محل‌یابی: باید محل جواب را در متن پیدا کرد.
۲	سلسله‌ای: باید چندین عمل محل‌یابی صورت گیرد.
۳	ترکیبی: اوّل باید چند محل جواب در متن پیدا شود، سپس ارتباط آن‌ها باهم شناخته شود.
۴	تولیدی: سؤالاتی که داوطلب برای پاسخ دادن به آن‌ها به تفسیر متن نیاز دارد.

توانمندی چهارم، پنجم و ششم موجود در چارچوب گوش‌کردن «تافل»، مربوط به دانش کاربردشناختی، دانش عمومی و نیز دانش نحوی در سطح گفتمان می‌شود. این نوع سؤالات عمدتاً از سطح سختی بالاتری برخوردار هستند زیرا علاوه بر سازه‌های کاربردشناختی و عمومی، دانش واژگانی، نحوی و آوایی افراد را نیز به‌صورت هم‌زمان می‌سنجند و مربوط به اهداف گوش‌کردن برای رسیدن به درک ابتدایی از متن و نیز گوش‌کردن برای یادگیری می‌شوند. پیوند اطّلاعات یک متن خواندن و یک متن گوش‌کردن نیز نام هفتمین توانمندی‌ای است که طراحان «تافل» برای سنجش مهارت گوش‌کردن در نظر گرفته‌اند. این نوع سؤالات توانمندی زبانی فراتر از سطح مهارت گوش‌کردن را می‌سنجند. علّت استفاده از چنین پرسش‌هایی ماهیّت ترکیبی تکالیف حقیقی زبانی است. مهارت‌های زبانی معمولاً به‌صورت مجزا اتّفاق نمی‌افتند به‌عنوان‌مثال، مهارت گوش‌کردن عمدتاً با صحبت کردن، خواندن و یا

نوشتن همراه است. قرار دادن این نوع سؤالات در آزمون به هدف گوش‌کردن برای ادغام اطّلاعات در بین چند متن مرتبط است.

چارچوب مشخصات تکلیف مهارت گوش‌کردن نیز نکات بسیار مهمی دربارهٔ سنجش مهارت گوش‌کردن در محیط دانشگاه در اختیار ما می‌گذارد. تکلیف‌های زبانی که در امتحان ساخته می‌شوند، باید، بازگوکننده‌ی محیط دانشگاهی باشند. بنا بر آنچه در بخش موقعیت چارچوب مشخصات تکلیف گفته شد این تکالیف می‌توانند شامل سخنرانی‌ها (در کلاس درس، در سالن کنفرانس)، مشاوره‌ها (گفتگوی استاد و دانشجو در اتاق استاد) و گفتگوها (گفت‌وگوی دو یا چند دانشجو، گفتگوی دانشجو و کارمندان بخش اداری دانشگاه) باشند.

همچنین متن سؤالات بخش گوش‌کردن باید دارای ویژگی‌های متون شنیداری دانشگاهی باشند:

۱- سرعت گفتار به‌طور میانگین ۳/۹۳ باشد؛
۲- ویژگی‌های زبان گفتار مانند مکث‌ها، پدیده‌ی سندی رعایت شوند؛
۳- عناصر انسجامی در متن بکارگرفته شود.

روش تحقیق

داده‌ها در این پژوهش از طریق برگزاری آزمون بسندگی زبان فارسی در مرکز بین‌المللی آموزش زبان فارسیِ دانشگاه فردوسی مشهد گردآوری گردیده است. این آزمون برای متقاضیان غیرفارسی‌زبان ورود به دانشگاه‌های ایران طرّاحی شده است و داوطلبان باید نمرهٔ حدّ نصاب (۶۰) را در آن کسب کنند. آزمون بسندگی زبان فارسی دانشگاه فردوسی مشهد، شامل چهار مهارتِ خواندن، شنیدن، شفاهی و نگارش و همچنین دستور و واژه می‌شود. در ادامه تنها به بررسی داده‌های به‌دست‌آمده از بخش گوش‌کردنِ یک نمونه از این آزمون که در تاریخ ۲۵ مرداد ۱۳۹۶ در مرکز بین‌المللی زبان فارسی برگزار گردید، پرداخته خواهد شد. این بخش از آزمون از دو متن تشکیل شده و شامل ۱۱ پرسش است. ویژگی سؤالات این بخش به تفصیل در جدول (۶) آمده است.

جدول (۵): مشخصات چارچوب گوش‌کردن (همان:۷۰)

مدلی برای سنجش مهارت گوش کردن در محیط دانشگاه		
توانایی‌های گوش کردن	**نوع تکلیف**	**نوع متن**
۱ـ درک جزئیات اصلی	تشخیص جواب پرسش‌های مربوط به چه کسی، چه چیزی، چه زمانی، چه مکانی	سخنرانی‌ها، مشاوره‌ها، گفت‌وگوها
۲- درک ایده‌ها، موضوع‌ها و اهداف اصلی متن	تشخیص و یا توضیح ایده‌های اصلی متن	
۳ـ درک واژگان کلیدی	تشخیص و درک واژگان کلیدی	
۴ـ درک کارکرد ارتباطی پاره‌گفتارها	درک کارکرد و یا هدف یک پاره‌گفتار	
۵- نتیجه‌گیری‌کردن برمبنای رابطه‌ی بین عناصر موجود در متن	تشخیص و توضیح رابطه‌ی بین ایده‌ها در متن	
۶ـ پیوند دادن اطّلاعات با دانش شخصی	ربط دادن اطّلاعات متن با دانش شخصی	
۷- پیوند دادن اطّلاعات یک متن شنیداری و یک متن خواندن	تحلیل و ترکیب اطّلاعات در بین چند متن	

شرکت‌کنندگان

تعداد ۹۰ نفر در این آزمون شرکت کردند که ۳۳ نفر آنها زن و ۵۷ نفر مرد بودند. عـراق بـا ۶۳ نفر بیشترین میزان شرکت‌کننده در این آزمون را داشـته اسـت و سـایر شـرکت‌کننـدگان از کشورهای اندونزی، لبنان، سوریه، استرالیا، کره‌ی‌جنوبی، پاکستان و یمن بوده‌اند. از نظـر رشـتهٔ تحصیلی، گروه مهندسی با ۳۸ نفر، پُرمخاطب‌ترین حوزهٔ دانشگاهی را تشکیل مـی‌دهـد. علـوم انسانی با ۱۷ نفـر، علـوم پایـه ۱۳ نفـر و پزشـکی ۷ نفـر در رتبـه‌هـای بعـدی از نظـر تعـداد شرکت‌کننده بوده‌اند. ۱۵ نفر از شرکت‌کنندگان نیز رشتهٔ تحصیلی‌شان نامشـخّص بـوده اسـت.

سطح تحصیلی برخی از شرکت‌کنندگان، کارشناسی و برخی کارشناسی ارشد و دکتری بوده است.

جدول (۶): سؤالات بخش گوش‌کردن آزمون بسندگی زبان فارسی دانشگاه فردوسی مشهد

شمارهٔ متن	نوع متن	سرعت گفتار	شرکت کنندگان	محیط	محتوا	شمارهٔ سؤال	نوع اطّلاعات خواسته‌شده	نوع مطابقت
۱	گفت‌وگوی تلفنی	۳/۹۲ هجا در ثانیه	مشتری ـ فروشنده	مکان عمومی		۱	تأیید	محل‌یابی
						۲	شیء	سلسله‌ای
						۳	صفت	محل‌یابی
						۴	محل	سلسله‌ای
						۵	مقدار	محل‌یابی
						۶	زمان	سلسله‌ای
۲	سخنرانی	۳/۲ هجا در ثانیه	استاد	محیط آموزشی		۷	وضعیت	سلسله‌ای
						۸	صفت	سلسله‌ای
						۹	صفت	محل‌یابی
						۱۰	عمل	محل‌یابی
						۱۱	شیء	محل‌یابی

جمع‌آوری داده‌ها

آزمون بسندگی زبان فارسی دانشگاه فردوسی مشهد در ساعت هشت روز ۲۵ مرداد ۱۳۹۶ در مرکز بین‌المللی زبان فارسی این دانشگاه در مدّت سه‌ساعت‌ونیم برگزار گردید. پاسخ‌گویی به سؤالات بخش گوش‌کردن در ساعت ۸:۱۵ آغاز شد و داوطلبان در مدّت زمان حدود ۲۰ دقیقه به ۱۱ سؤال آزمون پاسخ دادند. داوطلبان در ده کلاس مختلف بودند و در هر کلاس، حدود ۹ آزمون‌دهنده و یک مراقب حضور داشت. پس از آزمون، مدرّسان مجرب مرکز به کمک کلید آزمون که پیش‌تر طرّاحی شده بود، برگه‌ها را تصحیح کردند.

تجزیه‌وتحلیل داده‌ها

نرم‌افزار «وینستپس»[۱] (لیناکر، ۲۰۰۹)، ۱۱ پرسش بخش گوش‌کردن را توسّط مدل آماری «راش» تحلیل کرد. باید این نکته را در نظر داشت که استفاده از «مدل راش» برای تحلیل داده‌ها سه پیش‌شرط دارد: (همبلتون و همکاران،۱۹۹۱: ۱۲–۷) نخست اینکه داده‌های به دست آمده باید با مدل راش هم‌خوان باشد؛ به عبارتی، داده‌ها باید از میزان برازش قابل قبولی برخوردار باشند. ستون‌های «درون‌برازش»[۲] و «برون‌برازش»[۳]، برازش پرسش‌ها به مدل راش را نشان می‌دهند. مقادیر MNSQ باید بین ۰/۷۰ تا ۱/۳۰ باشند (باند و فاکس، ۲۰۰۸). برازش پرسش بدین معناست که پرسش با دیگر سؤال‌های آزمون هم‌ردیف است و به تعریف یک مقیاس برای سنجش سازهٔ موردنظر که در اینجا توانایی گوش‌کردن است، کمک می‌کند. از سوی دیگر آزمون باید تک‌بعدی باشد؛ یعنی تمامی پرسش‌ها توانایی واحدی را بسنجند (در اینجا مهارت گوش‌کردن). باید این نکته را در نظر داشت که در مهارت گوش‌کردن آزمون نمی‌تواند به‌طور کامل تک‌بعدی باشد زیرا مهارت شنیدن با سنجش سایر مهارت‌ها همراه است. به‌عنوان‌مثال، آزمون‌دهنده باید مطلبی را بنویسد (نوشتن) و یا باید مطلبی را بخواند و پاسخ دهد (مهارت خواندن). بااین‌حال، باید تنها یک مهارت برجسته و غالب در تمامی سؤالات مورد سنجش قرار گرفته باشد. جدول (۱) نشان می‌دهد که همهٔ پرسش‌ها، آمارهای درون‌برازش و برون‌برازش مناسبی دارند؛ بنابراین آزمون به میزان قابل قبولی تک‌بُعدی است.

شرط سوم به‌کارگیری مدل آماری «راش» این است که هر سؤال آزمون از دیگر سؤالات مستقل باشد؛ یعنی پاسخ‌گویی به یک پرسش تحت تأثیر پاسخ‌گویی به پرسش‌های دیگر نباشد. این نکته نیز در مورد سؤالات بخش گوش‌کردن آزمون زبان فردوسی صدق می‌کند. نتیجه بررسی نشان داد که پایایی آزمون بر اساس «راش» ۰/۶۶ و بر مبنای مدل «آلفای کرونباخ» ۰/۷۳ است. این در حالی است که برای سنجش مهارت گوش‌کردن، پایایی آزمون بهتر است بین ۰/۸۰ تا ۰/۸۹ باشد (لادو،۱۹۶۱ نقل در هیوز،۳۲:۱۹۸۱).

---

۱. WINSTEPS

۲. Infit

۳. Outfit

جدول (۷): اندازهٔ دشواری پرسش‌ها و برازش آن‌ها به «مدل راش».

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	ITEM	G
1	68	90	.11	.28	.96	-.3	1.01	.1	.46	.44	76.7	78.3	L1	D
2	156	90	-.27	.22	.95	-.1	.55	-.8	.57	.53	82.6	81.2	L2	0
3	166	90	-.64	.26	1.43	1.2	2.47	1.2	.36	.49	89.5	90.9	L3	0
4	176	90	-2.16	.48	.75	-.3	1.05	.5	.31	.31	96.5	96.3	L4	0
5	93	90	1.49	.14	1.02	.2	.92	-.1	.64	.64	60.5	56.5	L5	0
6	126	90	.27	.20	.97	-.2	.99	.0	.57	.56	58.1	61.8	L6	0
7	166	90	-.60	.32	.97	-.1	1.17	.6	.39	.40	87.2	85.7	L7	0
8	85	90	1.69	.16	1.03	.3	.96	-.2	.60	.61	48.8	52.5	L8	0
9	139	90	.39	.17	.88	-.6	.63	-.8	.63	.59	67.4	68.2	L9	0
10	103	90	1.26	.15	.99	.0	.90	-.3	.63	.63	52.3	52.6	L10	0

نقشهٔ آزمون‌دهنده ــ پرسش

شکل (۱) «نقشهٔ رایت»[۱] یا نقشهٔ آزمون‌دهنده ـ پرسش است. در این نقشه، سطح توانایی زبانی آزمون‌دهندگان و سطح دشواری سؤالات به‌طور هم‌زمان نمایش داده شده است و درنتیجه مستقیماً قابل‌مقایسه هستند و همان‌گونه که پیش‌تر گفته شد این امکان مقایسه هم‌زمان، از جمله مزیّت‌های مدل «راش» است که در مدل‌های کلاسیک وجود ندارد. در سمت راست نقشه، سطح سختی پرسش‌ها و در سمت چپ سطح توانایی آزمون‌دهندگان قرار دارند. انتظار می‌رود که سؤالات آزمون متناسب با سطوحِ زبانیِ مختلفِ آزمون‌دهندگان طراحی شده باشند تا توانایی تمییز داوطلبان ضعیف، متوسط و قوی را داشته باشند. به عبارتی، توده‌ی سؤالات در سمت راست باید مقابل توده‌ی آزمون‌دهندگان در سمت چپ قرار بگیرند. اگر در نقشه‌ای در مقابل پرسش و مجموعه‌ای از پرسش‌ها، هیچ داوطلبی قرار نگیرد و یا تعداد کمی داوطلب دیده شود؛ این نکته نشان می‌دهد که آن پرسش یا آن مجموعه‌ی پرسش‌ها کارکرد چندانی در تمییز توانایی داوطلبان از یکدیگر ندارند و باید مورد بازبینی قرار بگیرند.

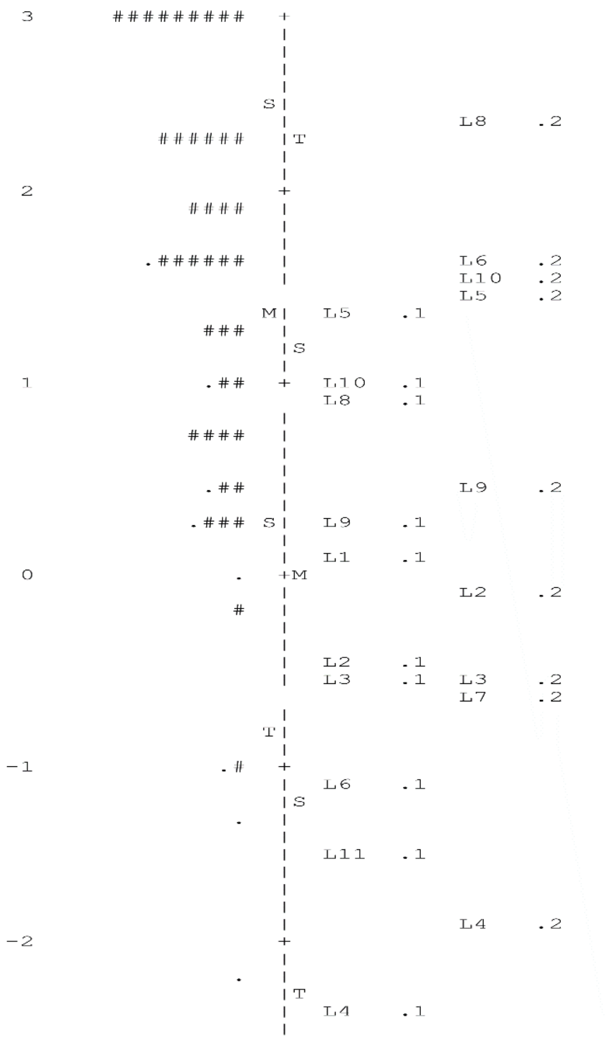
همان‌طور که دیده می‌شود، ۱۹ آزمودنی در بالای مقیاس قرار دارند که در سطح بالای توانایی هستند، امّا تنها یک پرسش در آن سطح وجود دارد؛ بنابراین آزمون برای تشخیص

۱. Wright map

آزمون‌دهندگان قوی ناکارآمد است. نقشه همچنین نشان می‌دهد که پرسش‌های آسـان بسـیاری در انتهای مقیاس قرار دارند که در آن سطح تعداد آزمودنی کمی قـرار دارد؛ بنـابراین بـه دلیـل سادگی بسیار زیاد، این پرسش‌ها برای تمییز آزمون‌دهندگان ضعیف از متوسط، چنـدان کارآمـد نیستند.

شکل (۱): نقشهٔ آزمون‌دهنده-پرسش

سطح سؤالات آزمونسطح آزمون‌دهندگان



بحث و نتیجه‌گیری

مدل آماری «راش» نشان داد که پایایی بخش گوش‌کردن آزمون زبان فارسی فردوسی ۰/۶۶ است که برای یک آزمون سرنوشت‌ساز میزان مناسبی به‌شمارنمی‌آید. لازم به یادآوری است که پایایی بخش گوش‌کردن آزمون تافل ۰/۸۷ است (اجوکیشنال تستینگ سیستم، ۲۰۱۸: ۷). علّت پایین بودن پایایی آزمون زبان فردوسی را می‌توان در تعداد کم سؤالات و تکلیف‌های آزمون دانست. این آزمون تنها از دو تکلیف و یازده پرسش تشکیل شده است که در مقایسه با آزمون بسندگی «تافل» که تعداد تکلیف‌های زبانی آن حدود ۶ عدد و تعداد سؤالات آن بیش از ۴۰ مورد است (اجوکیشنال تستینگ سیستم، ۲۰۰۹: ۶) رقم بسیار پایینی به‌حساب می‌آید. از منظر سازه‌ها می‌توان گفت که هیچ‌کدام از پرسش‌های آزمون زبان فارسی فردوسی به سنجش دانش نحوی در سطح گفتمان و یا دانش کاربردشناختی آزمون‌دهندگان نمی‌پردازند. تمرکز این آزمون بیشتر بر سنجش سازه‌های آوایی و واژگانی است. از همین رو آزمون زبان فارسی فردوسی تصویر ناقصی از توانمندی گوش‌کردن داوطلبان ارائه می‌دهد.

از سوی دیگر مطابقت آزمون با چارچوب مشخصات تکلیف نکات مهم دیگری را روشن می‌سازد. تکلیف شمارهٔ یک آزمون زبان فارسی فردوسی گفت‌وگویی تلفنی بین یک پلیس و یک مراجعه‌کننده است و تکلیف شمارهٔ دو نیز شامل یک سخنرانی علمی می‌شود. این در حالی است که چارچوب مشخصات تکلیف در بخش موقعیتِ تکلیف‌هایِ مهارت گوش‌کردن، شرکت‌کنندگان در تکالیف زبانی را استاد، کتابدار، مسئول ثبت‌نام، منشی گروه، راهنمای دانشجویان بین‌الملل، مسئول آزمایشگاه، دستیار استاد، پرستار و یا دکتر در مرکز درمانی دانشگاه معرفی می‌کند؛ بنابراین گفتگوی بین پلیس و مراجعه‌کننده نمی‌تواند نمونهٔ مناسبی از فعّالیّتهای زبانی مربوط به مهارت گوش‌کردن در محیط دانشگاه باشد.

همچنین در چارچوب مشخصات تکلیف چهار هدف برای فعّالیّت‌های گوش‌کردن در محیط دانشگاه مشخص شد: ۱ـ پیدا کردن اطّلاعات ۲ـ رسیدن به درک ابتدایی از متن ۳ـ یادگیری ۴ـ ادغام اطّلاعات در بین چند متن مختلف؛ اما آنچه در آزمون زبان فارسی فردوسی مورد سنجش قرار می‌گیرد تنها گوش‌کردن با هدف پیدا کردن اطّلاعات است؛ بنابراین سایر

اهداف مهارت گوش‌کردن در این آزمون سنجیده نمی‌شوند که خود ضعفی جدّی به شمار می‌آید.

بااین‌حال از جمله ویژگی‌های زبانی مثبت تکلیف‌های آزمون زبان فارسی فردوسی، می‌توان به سرعت گفتار در آنها اشاره کرد. سرعت گفتار متن نخست تقریباً برابر با میانگین سرعت واقعی گفتار در زبان فارسی یعنی ۳/۹۳ است اما متن دوم از سرعت پایینی برخوردار ۳/۲ است و این امر ساده شدن این بخش از آزمون را به همراه داشته است. هرچند طراحان آزمون تلاش کرده‌اند تا ویژگی‌های زبان گفتاری مانند تکیه، آهنگ کلام، سرعت گفتار و پدیده‌ی سندی را در فایل صوتی آزمون رعایت کنند اما تقریباً در هیچ بخشی از آزمون مکث‌های طبیعی گفتار مانند واژه‌های مکث‌پُرکن، حروف کشیده، شروع‌های اشتباهیِ ردیابی شده و نیز تکرار واجی دیده نمی‌شود.

از سوی دیگر نمودار آزمون‌دهنده – پرسش نشان‌داد که آزمون بیشتر در سطح آزمون‌دهندگان بسیار ضعیف، ضعیف و متوسط است. این آزمون توانایی تمییز داوطلبان قوی از متوسط را ندارد و حتی برخی از پرسش‌های آن به‌اندازه‌ای ساده طراحی شده‌اند که از سطح تمامی آزمون‌دهندگان پایین‌ترند. علّت این امر را می‌توان در نوع سؤالات آزمون یافت. از نظر نوع اطّلاعات خواسته شده، پرسش‌های آزمون عمدتاً شامل تشخیصِ محل، مکان، شیء، زمان، مقدار و یا صفت می‌شود که عمدتاً در سطوح سختی پایینی هستند. در مقابل در آزمون‌هایی مانند تافل، سؤالات بسیاری مربوط به تشخیص تفاوت‌ها، معادل‌یابی‌ها برای واژگان کم‌بسامد، روابط علت‌ومعلولی و شباهت‌ها می‌شود که دارای سطح سختی بالاتری هستند. طبیعتاً آزمون به سؤالات بیشتری از این نوع نیازمند است. از نظر نوع مطابقت نیز، تمامی سؤالات، محل‌یابی و سلسله‌ای هستند. طراحان می‌توانند با به‌کارگیری سؤالاتی از نوع مطابقت‌های ترکیبی و تولیدی که سطح سختی بالاتری دارند دشواری آزمون را افزایش دهند تا آزمون بتواند توانایی زبانی آزمون‌دهندگان با سطوح پیشرفته‌ی زبانی را نیز بسنجد.

کتاب‌نامه

۱- جلیلی، س.ا. (۱۳۹۰). آزمون مهارتی فارسی (آمفا) بر پایۀ چهار مهارت اصلی زبانی. *پایان‌نامۀ کارشناسی‌ارشد*. دانشکدۀ ادبیّات فارسی و زبان‌های خارجی دانشگاه علاّمه طباطبائی، تهران، ایران.

۲- عرفانیان قونسولی، ل.، و شریفی، ش.، و تقی‌پور، م. (۱۳۹۳). بررسی الگوهای مکث در زبان فارسی. *مجموعه‌ی مقاله‌های سومین همایش تحلیل گفتمان و کاربردشناسی*. تهران: نشر نویسهٔ پارسی.

۳- قنسولی، ب. (۱۳۸۹). طرّاحی و روایی‌سنجی آزمون بسندگی زبان فارسی. *پژوهش زبان‌های خارجی*، ۵۷، ۱۱۵_ ۱۲۹.

۴- گلپور، ل. (۱۳۹۴). طرّاحی و اعتباربخشی آزمون بسندگی زبان فارسی بر پایهٔ چهار مهارت زبانی. *رسالهٔ دکتری*. دانشکدهٔ ادبیّات فارسی و زبان‌های خارجی دانشگاه پیام‌نور مرکز، تهران، ایران.

۵- مخلصین، م.، و احمدی‌زاده، ز.، و کسبی، ف.، و احدی، ح. (۱۳۹۶). تأثیر سن، جنسیت و تکلیف بر سرعت گفتار فارسی‌زبانان. *فصلنامه‌ی کومش*، ۱۹(۲)، ۳۲۷_۳۳۲.

۶- موسوی، ث. (۱۳۷۹). طرّاحی و اعتباربخشی آزمون استاندارد زبان فارسی بر پایهٔ چهار مهارت زبانی. *پایان‌نامهٔ کارشناسی‌ارشد*. دانشکدهٔ ادبیّات و علوم انسانی دانشگاه شیراز، شیراز، ایران.

۷- وکیلی‌فرد، ا.ر.، و کولی‌وندی، ا.، و فاطمی‌منش، ع. (۱۳۹۲). بسامد تکرار زمان‌های دستوری در گفتار و نوشتار فارسی: در جست‌وجوی اولویت‌های آموزشی دستور برای فارسی‌آموزان. *پژوهش‌نامه‌ی آموزش زبان فارسی برای غیرفارسی‌زبانان*، ۲(۳)، ۹۹_۱۱۵.

۸- Bachman, L. (۲۰۰٤). *Statistical Analysis for Language Assessment*. New York: Cambridge University Press.

۹- Bachman, L., & Palmer, A. (۲۰۱۰). *Language Assessment in practice*. New York: Oxford University Press.

۱۰- Bejar, I., & Douglas, D., & Jamieson, J., & Nissan, S., & Turner, J. (۲۰۰۰). *TOEFL ۲۰۰۰ Listening Framework: A working paper* (TOEFL Monograph No. ۱۹) Princeton, NJ: Educational Testing Service.

۱۱- Biber, D. (۱۹۸۸). *Dimensions of register variation*. New York: Cambridge University Press.

۱۲- Bond, T. G., & Fox, C. M. (۲۰۰۷). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum.

۱۳- Buck, G. (۲۰۰۱). *Assessing Listening. Cambridge*: Cambridge University Press.

۱۴- Carrol, J. B. (۱۹۹۲*). Human Cognative Abilities: A Survey of Factor Analysis*. Cambridge: Cambridge University Press

۱۵- Chapelle, C. A., & Enright M. K., & Jamieson, J. M. *Building a Validity Argument for the Test of English as a Foreign Language.* New York: Routledge.

۱۶- Educational Testing System, (۲۰۰۹). *The official Guide to the new TOEFL*. New York: McGraw-Hill.

۱۷- Educational Testing System, (۲۰۱۸). Reliability and Comparability of TOEFL iBT Score. *https://www.ets.org/s/toefl/pdf/toefl_ibt_research_s ۱۷۳.pdf*.

۱۸- Jamieson, J., & Jones, S., & Kirsch, I., & Mosenthal, P., & Taylor, C. (۲۰۰۰). *TOEFL ۲۰۰۰ Framework: A Working Paper* (TOEFL monograph No. ۱۶). Princeton, NJ: Educational Testing system.

۱۹- Jamieson, J. M., & Eigner, D., & Grabe W., & Kunnan, A. J. (۲۰۰۸). "*Frameworks for a new TOEFL*".In Building a Validity Argument for the Test of English as a Foreign Language Edited by Chapelle, C. A., Enright M. K., Jamieson, J. M. New York: Routledge. pp. ۱-۲۵.

۲۰- Hambleton, R. K., & Swaminathan, H., & Rogers, J. (۱۹۹۱*). Fundamentals of Iten Response Theory*. Newbury Park: Sage Publication.

۲۱- Hughes, A. (۱۹۸۹). *Testing for Language Teaching*. New York: Cambridge University Press.

۲۲- Lado, R. (۱۹٦۱). *Language Testing*. London: Longman.

۲۳- Linacre, J. M. (۲۰۰۹). *WINSTEPS Rasch Measurement* [Computer program]. Chicago, IL: Winsteps.

۲۴- Messick, S. (۱۹۹۴). The interplay of evidence and Consequences in the Validation of Performance Assessment. *Educational Researcher*, ۲۳(۲), ۱۳-۲۳.

۲۵- Miller, c., & Aghajanian-Stewart, K. (۲۰۱۸). *A Frequency Dictionary of Persian*. New York: Routledge.

۲۶- Mousavi, S. A. (۲۰۱۲). Item Response Theory. In *An Encyclopedic Dictionary of Language Testing*. ۵th ed. Tehran. Rahnama.

۲۷- Rost, M. (۲۰۱۶). *Teaching and Researching Listening*. New York: Routledge.

۲۸- Powers, D. (۱۹۸۶). Academic demands related to listening skills. *Language Testing*, ۳, ۱-۳۸.

۲۹- Spolsky, B. (۲۰۰۸). Language Assessment in Historical and future Perspective. In Hornberger, Nancy (Ed.). *Encyclopedia of Language and Education*. ۲nd ed.Vol ۷. New York: Springer.